

Status report on a different method of performing the ND Slicing

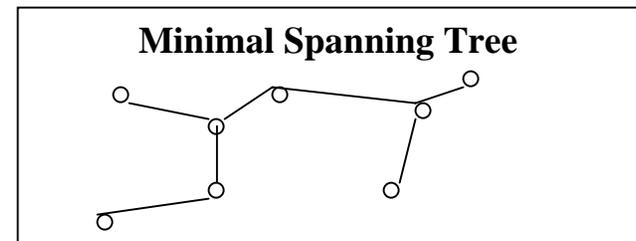
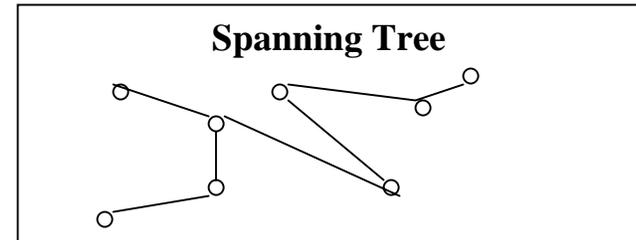
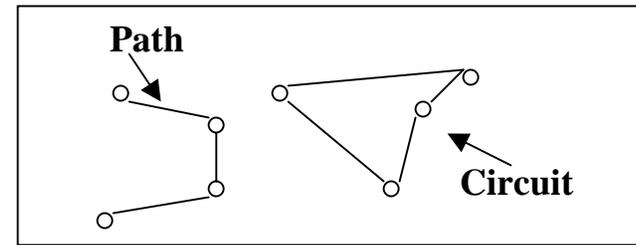
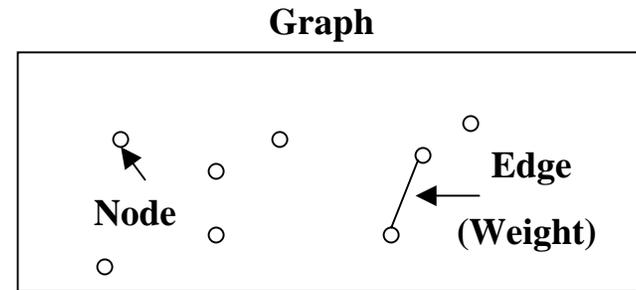
N. Saoulidou, Fermilab, Collaboration Mtg, 09-16-04

Outline

- Description of the clustering technique (MST)
 - MST definition
 - Properties
 - MST & Clustering
 - Metric for weight definition
- Very first results & comparison with current SR code :
 - Slice purity & completeness
 - Track purity & completeness
 - Shower purity & completeness
- Summary - On going work

-Minimal Spanning Trees Basics-

- An **edge weighted Linear Graph** is composed of a set of points called **nodes** and a set of node pairs called **edges** with a number called **weight** assigned to each edge.



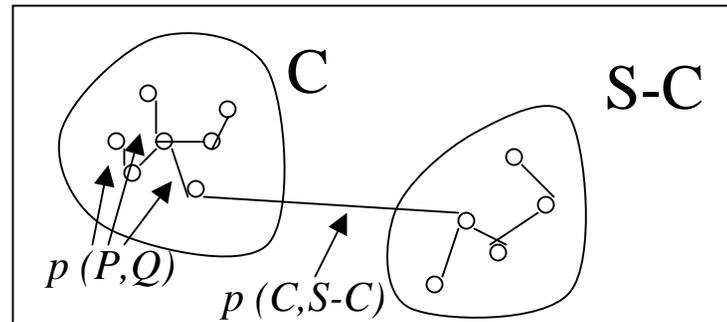
- A **path** in a graph is a sequence of edges joining two nodes. A **circuit** is a closed path.

- A **spanning tree** is a connected graph with no circuits which contains all nodes.

- A **minimal spanning tree** is the spanning tree whose weight (= sum of weights of its constituent edges) is minimum among all spanning trees in this set of nodes.

- MST Theorem 1 -

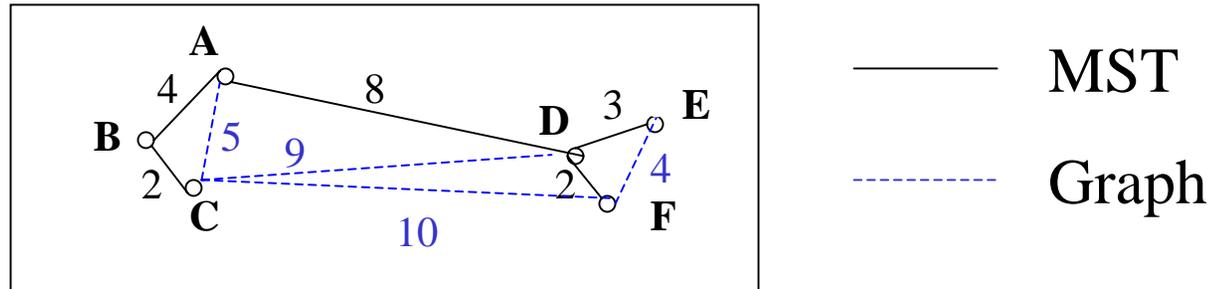
- Theorem 1 : " If S denotes the nodes of G and C is a nonempty subset of S with the property that $p(P,Q) < p(C,S-C)$ for all partitions (P,Q) of C then the restriction of any MST to the nodes of C forms a connected subtree of the MST ". The significance of this theorem for cluster detection can be illustrated if the following figure which depicts the MST for a point set containing two clusters C and $S-C$:



- This theorem assures us that **the subgraph of an MST does not break up the real clusters in S** , but on the other hand neither does it force breaks where real gaps exist in the geometry of the point set.
- A spanning tree is forced by its very nature to span all the points but at least **the MST jumps across the smaller gaps first**.

- MST Theorem 2-

- Theorem 2 : "If T is an MST for graph G and X, Y are two nodes of G , then the unique path in T from X to Y is a minimax path from X to Y ".[1]



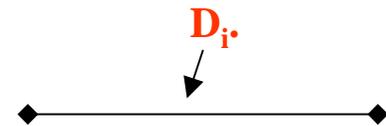
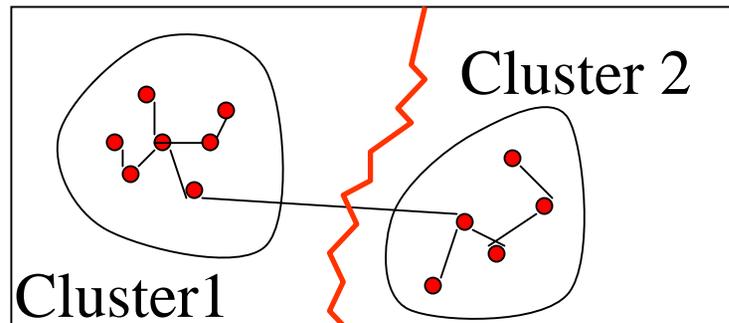
- **Cost** : maximum edge weight of the path e.g the path (CBADE) has a cost of 8.
- **Minimax path** : The path between a pair of nodes that has the least cost e.g there are four minimax paths from C to F all of cost 8.
- The **minimax path** each of whose **subpaths** are also **minimax** lies **within the MST** and that is not a coincidence as shown in the previous theorem.
- So the preference of minimax paths in the **MST** forces it to **connect** two **nodes X and Y** belonging to a **tight cluster** without straying outside the cluster.

- MST properties -

- **The MST is deterministic.** It does not depend on random choices in the algorithm or on the order in which nodes and edges are selected and examined but only on the given set of nodes.
- **The MST is invariant under similarity transformations,** that is under all transformations that preserve the monotony of the metric (rotations, translations changes of the scale and even some nonlinear distortions).
- **The metric for the weight assignment** can be defined in many ways and does not have to be the **Euclidean Distance between 2 nodes**.

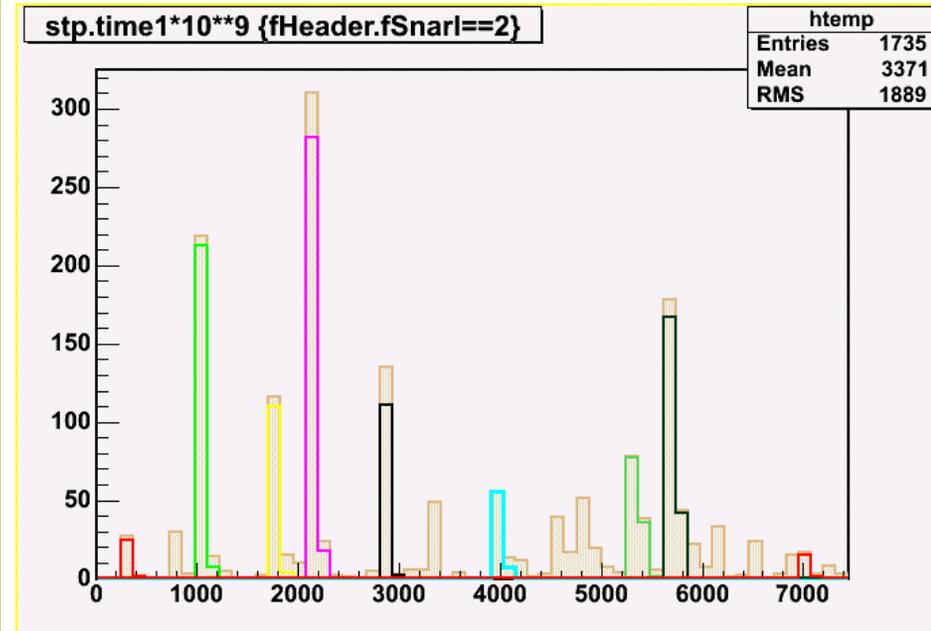
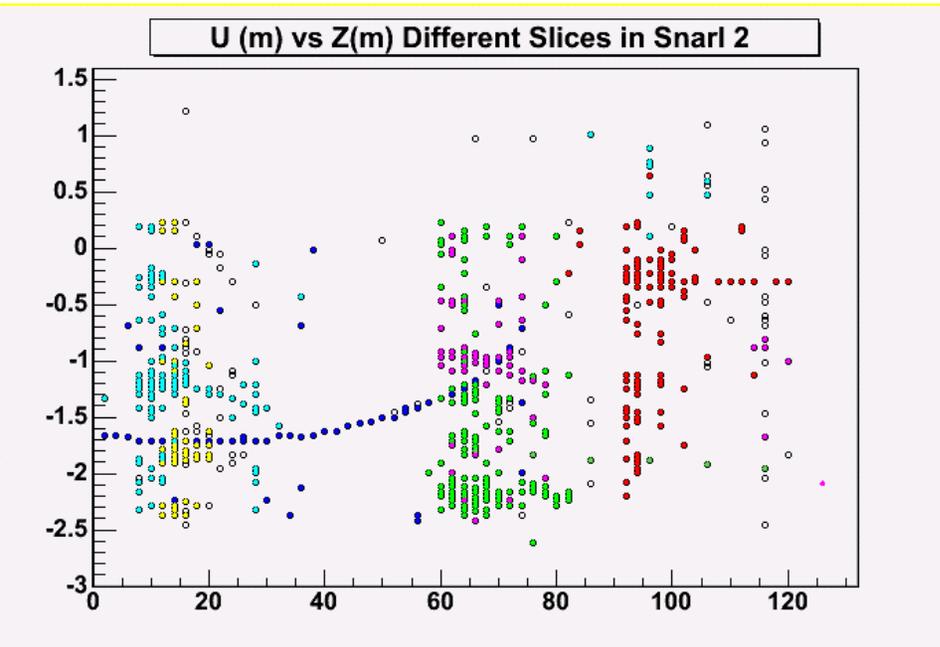
-MST & Cluster Analysis-

- Main Idea : *After forming the MST of a set of points group the points into disjoint sets by joining all edges of weights D_i or less. Each set is then said to form a cluster at level D_i . Thus all segments joining two clusters defined at level D_i will have lengths greater than D_i .*



MST in ND (Clustering in time & space)

- I have used a hybrid metric for constructing the length of the MST branch connecting two strips :
 - time difference $\times c + a \times$ length difference (in z) and I using a semi-random initial cut that is going to be tuned...
 - This method has just ONE parameter that needs to be tuned...

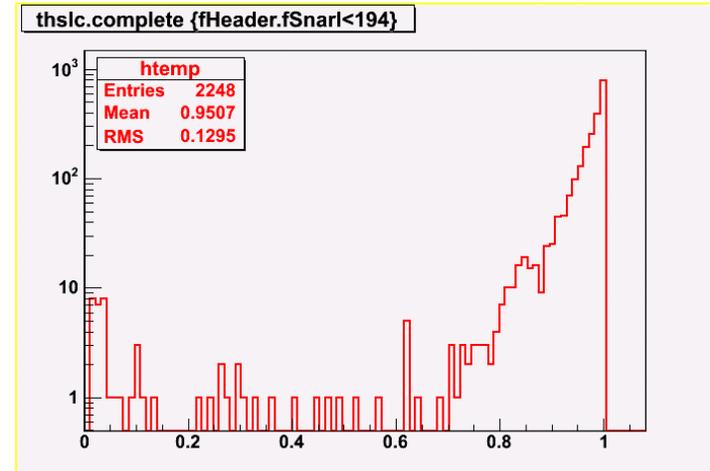
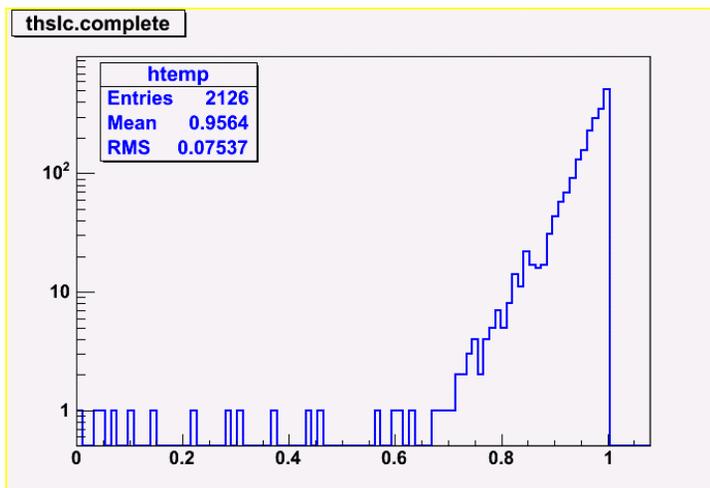
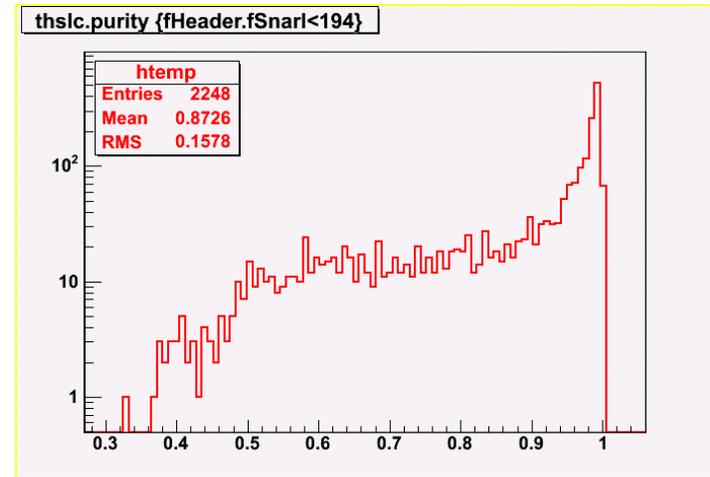
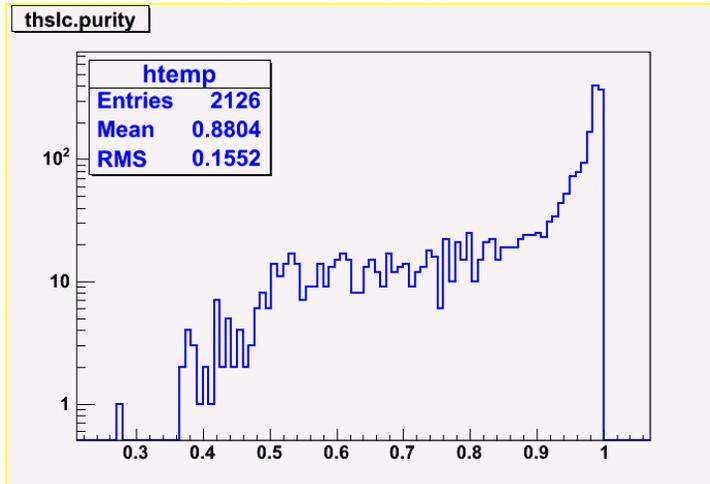


MST results & comparison with SR

NEW

Slice purity (top) & completeness (bottom)

SR



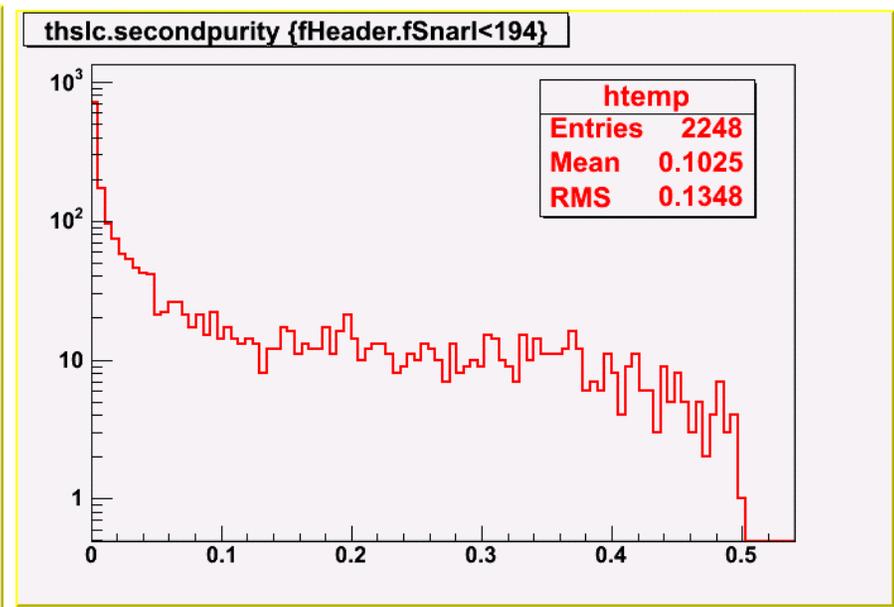
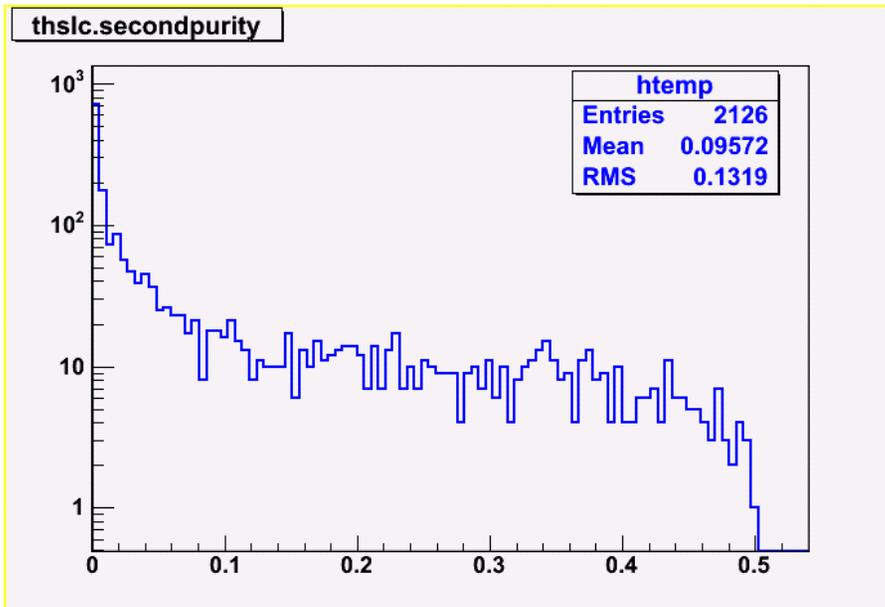
- The MST is as good as the SR...

MST results & comparison with SR con't

NEW

Slice contamination from a
second neutrino event

SR



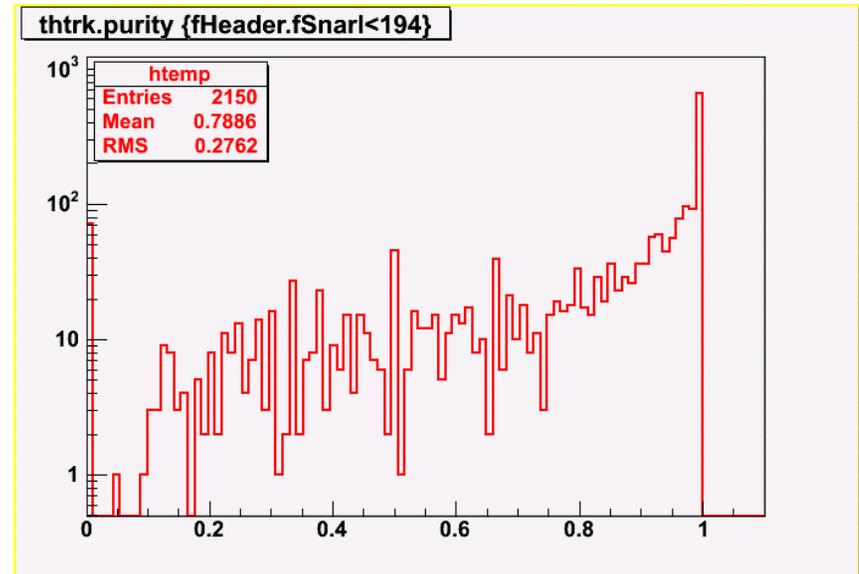
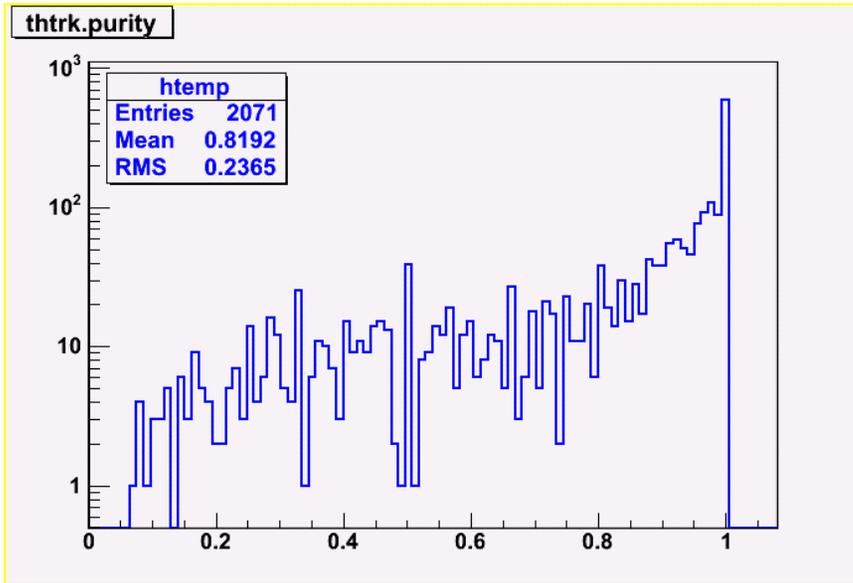
- The MST is as good as the SR...

MST results & comparison with SR : tracking

Track purity

NEW

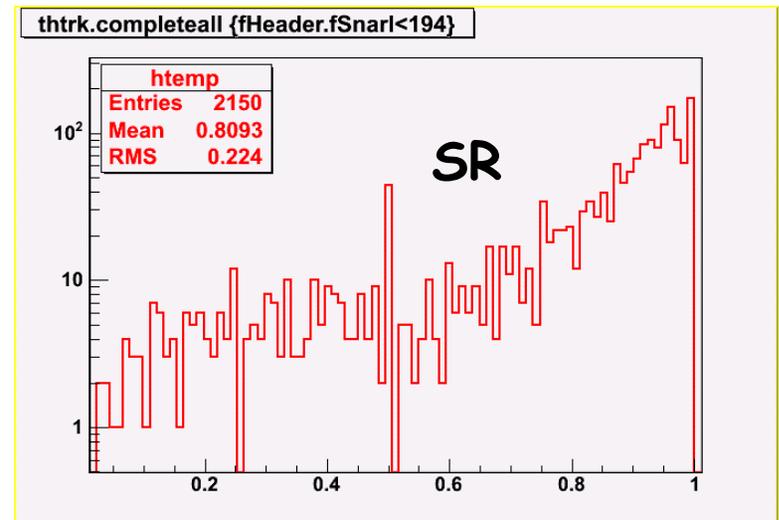
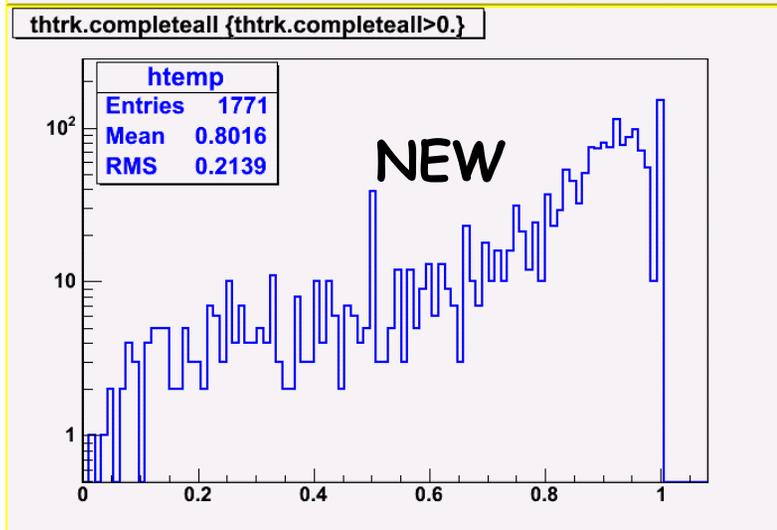
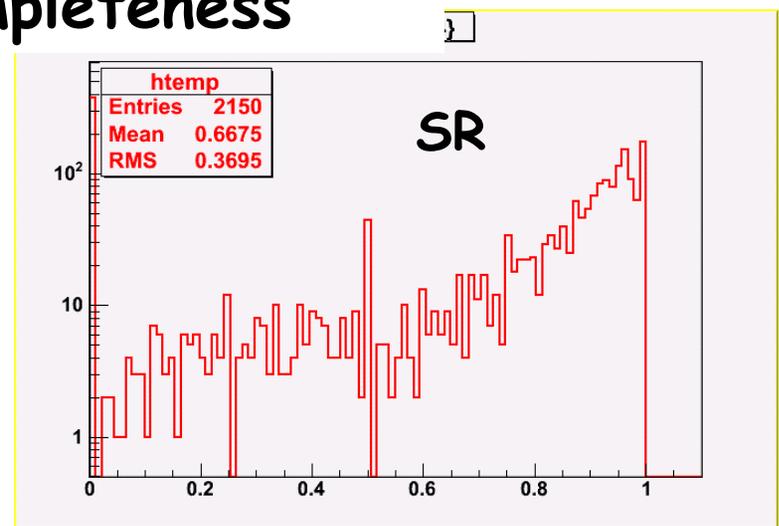
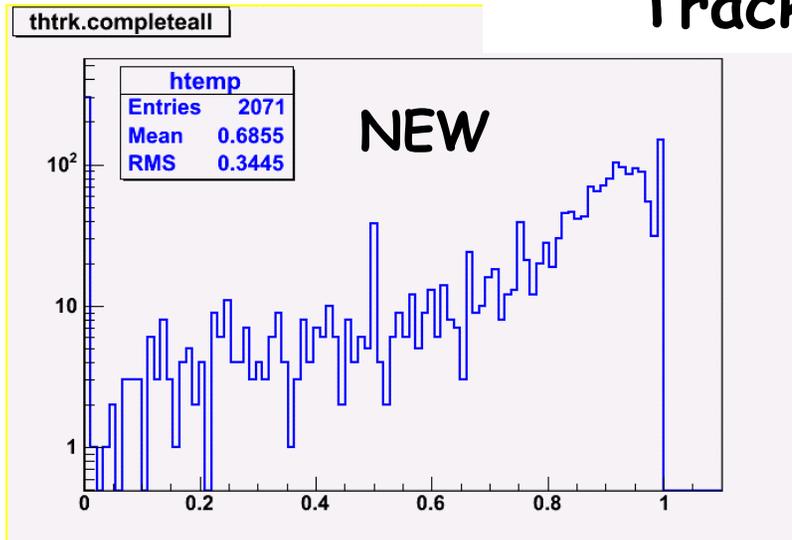
SR



- The MST is better than the SR (given that the changes that I see are not due to the fact that these files run with ~ 4 days difference).

MST results & comparison with SR : tracking

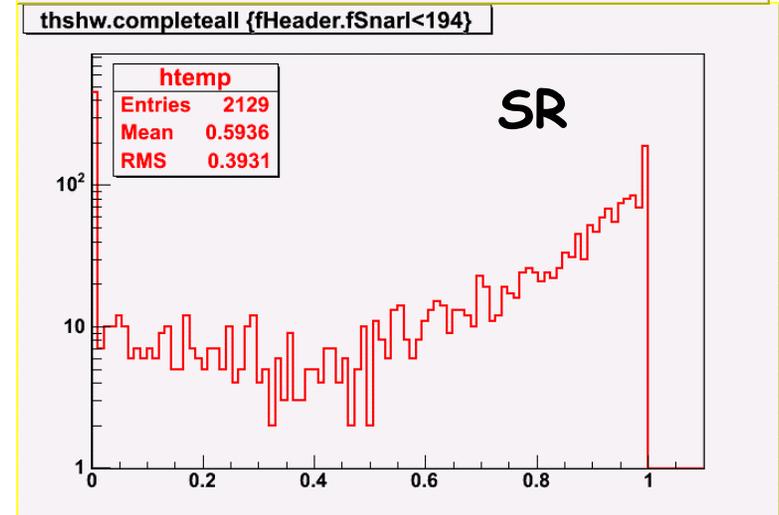
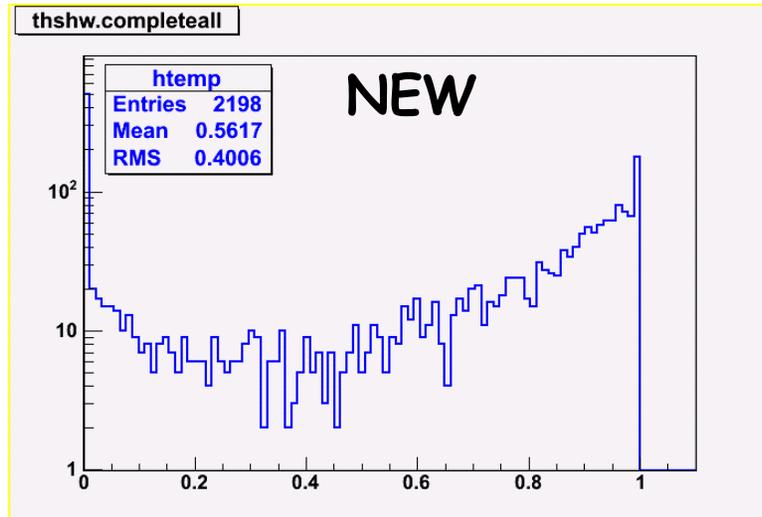
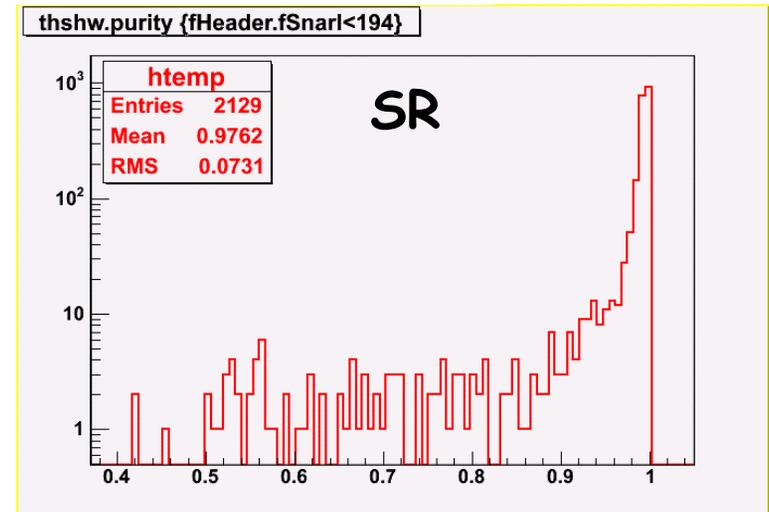
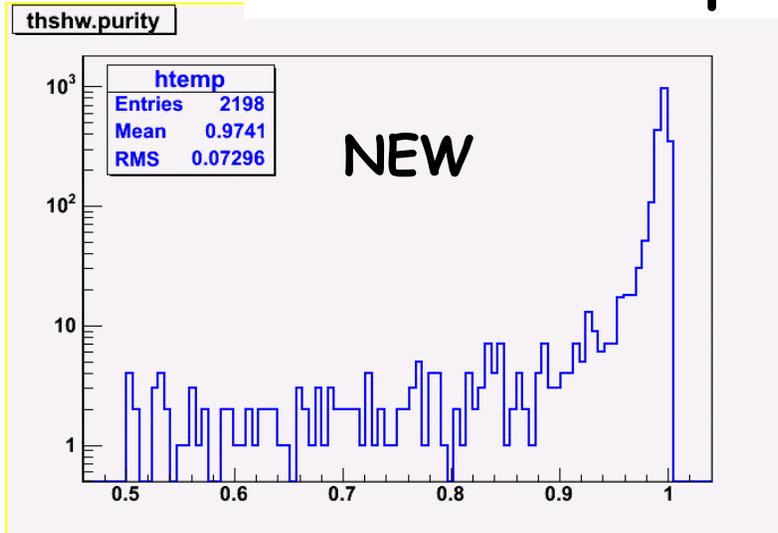
Track completeness



- The MST as good or slightly better than SR

MST results comparison with SR : Shower finding

Shower completeness & purity



- The MST is as good as the SR ...(slightly worse on shower completeness)

Summary & Ongoing work

- This new slicing method seems quite promising (initial results very very close to SR)
- It is flexible due to the freedom one has on the metric definition for the weight assignment
- It can be used to either strips or digits...
- I plan to start tuning the cut and also look into more detail on timing ND issues in order to further improve (if possible) the slicing.